

How can I select specific columns in R, and what are some examples of doing so?

Authored by
stats writer

May 2, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I select specific columns in R, and what are some examples of doing so?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=141854>

In R, specific columns can be selected from a data frame using the "select" function from the dplyr package. This function allows users to specify the columns they want to keep or remove from a data frame, based on their names or indices. For example, to select columns "species" and "height" from a data frame named "plants", the code would be: `select(plants, species, height)`. Another way to select columns is by using the "subset" function, which allows users to specify logical conditions to filter the data frame and select only the desired columns. For instance, `subset(plants, height > 5)` would select all columns from the "plants" data frame where the height is greater than 5. These functions provide users with flexibility in selecting specific columns, making data manipulation and analysis more efficient.

Select Specific Columns in R (With Examples)

You can use the following syntax to select specific columns in a data frame in base R:

#select columns by name

df

#select columns by index

df

Alternatively, you can use the `select()` function from the package:

library(dplyr)

#select columns by name

df %>%

select(col1, col2, col4)

#select columns by index

```
df %>%
```

```
select(1, 2, 4)
```

For extremely large datasets, it's recommended to use the `dplyr` method since the `select()` function tends to be quicker than functions in base R.

The following examples show how to use both of these methods in practice with the following data frame:

```
#create data frame
```

```
df <- data.frame(a=c(1, 3, 4, 6, 8, 9),
```

```
b=c(7, 8, 8, 7, 13, 16),
```

```
c=c(11, 13, 13, 18, 19, 22),
```

```
d=c(12, 16, 18, 22, 29, 38))
```

```
#view data frame
```

```
df
```

```
a b c d
```

```
1 1 7 11 12
```

```
2 3 8 13 16
```

```
3 4 8 13 18
```

```
4 6 7 18 22
```

```
5 8 13 19 29
6 9 16 22 38
```

Example 1: Select Specific Columns Using Base R (by name)

The following code shows how to select specific columns by name using base R:

```
#select columns by name
df
```

```
a b d
1 1 7 12
2 3 8 16
3 4 8 18
4 6 7 22
5 8 13 29
6 9 16 38
```

Example 2: Select Specific Columns Using Base R (by index)

The following code shows how to select specific columns by index using base R:

```
#select columns by index
df
```

```
a b d
1 1 7 12
2 3 8 16
3 4 8 18
4 6 7 22
5 8 13 29
6 9 16 38
```

Example 3: Select Specific Columns Using dplyr (by name)

The following code shows how to select specific columns by name using dplyr:

```
library(dplyr)
```

```
#select columns by name
```

```
df %>%
```

```
select(a, b, d)
```

```
a b d
1 1 7 12
2 3 8 16
3 4 8 18
4 6 7 22
5 8 13 29
```

6 9 16 38

Example 4: Select Specific Columns Using dplyr (by index)

```
library(dplyr)
```

```
#select columns by index
```

```
df %>%
```

```
select(1, 2, 4)
```

```
a b d
```

```
1 1 7 12
```

```
2 3 8 16
```

```
3 4 8 18
```

```
4 6 7 22
```

```
5 8 13 29
```

```
6 9 16 38
```